ET&C
FOCUS

Focus articles are part of a regular series intended to sharpen understanding of current and emerging topics of interest to the scientific community.

An International Perspective on the Tools and Concepts for Effluent Toxicity Assessments in the Context of Animal Alternatives: Reduction in Vertebrate Use

Teresa J. Norberg-King,[a] Michelle R. Embry,[b,*] Scott E. Belanger,[c] Thomas Braunbeck,[d] Joshua D. Butler,[e] Phil B. Dorn,[f] Brianna Farr,[b] Patrick D. Guiney,[g] Sarah A. Hughes,[h] Marlo Jeffries,[i] Romain Journel,[j] Marc Lèonard,[k] Mark McMaster,[l] James T. Oris,[m] Kathy Ryder,[n] Helmut Segner,[o] Thomas Senac,[j] Glen Van Der Kraak,[p] Graham Whale,[q] and Peter Wilson[r]

[a]US Environmental Protection Agency, Duluth, Minnesota
[b]Health and Environmental Science Institute, Washington, DC, USA
[c]The Procter & Gamble Company, Cincinnati, Ohio, USA
[d]University of Heidelberg, Heidelberg, Germany
[e]ExxonMobil, Annandale, New Jersey, USA
[f]PB Dorn Associates, Galveston, Texas, USA
[g]University of Wisconsin, Madison, Wisconsin, USA
[h]Shell Health–Americas, Houston, Texas, USA
[i]Texas Christian University, Fort Worth, Texas, USA
[j]Sanofi, Paris, France
[k]L'Oréal, Aulnay-sous-Bois, France
[l]Environment and Climate Change Canada, Burlington, Ontario
[m]Miami University, Oxford, Ohio, USA
[n]UK Home Office, Dundee, United Kingdom
[o]University of Bern, Bern, Switzerland
[p]University Guelph, Guelph, Ontario, Canada
[q]Shell International BV, Manchester, United Kingdom
[r]Sanofi, Bridgewater, New Jersey, USA

Abstract: Since the 1940s, effluent toxicity testing has been used to assess potential ecological impacts of effluents and help determine necessary treatment options for environmental protection prior to release. Strategic combinations of toxicity tests, analytical tools, and biological monitoring have been developed. Because the number of vertebrates utilized in effluent testing is thought to be much greater than that used for individual chemical testing, there is a new need to develop strategies to reduce the numbers of vertebrates (i.e., fish) used. This need will become more critical as developing nations begin to use vertebrates in toxicity tests to assess effluent quality. A workshop was held to 1) assess the state of science in effluent toxicity testing globally; 2) determine current practices of regulators, industry, private laboratories, and academia; and 3) explore alternatives to vertebrate (fish) testing options and the inclusion of modified/new methods and approaches in the regulatory environment. No single approach was identified, because of a range of factors including regulatory concerns, validity criteria, and wider acceptability of alternatives. However, a suite of strategies in a weight-of-evidence approach would provide the flexibility to meet the needs of the environment, regulators, and the regulated community; and this "toolbox" approach would also support reduced reliance on in vivo fish tests. The present Focus article provides a brief overview of wastewater regulation and effluent testing approaches. Alternative methodologies under development and some of the limitations and barriers to regulatory approaches that can be selected to suit individual country and regional requirements are described and discussed. *Environ Toxicol Chem* 2018;37:2745–2757. © 2018 The Authors. *Environmental Toxicology and Chemistry* published by Wiley Periodicals, Inc. on behalf of SETAC.

Keywords: Effluent testing; Animal alternatives; Regulation; Weight of evidence; Toxicity testing; Wastewater

* Address correspondence to: membry@hesiglobal.org

Published online in Wiley Online Library (wileyonlinelibrary.com).
DOI: 10.1002/etc.4259

INTRODUCTION

Many thousands of chemicals continue to be released each year into the environment via permitted and most often regulated effluent discharges (Figure 1). Chemical-by-chemical controls are used for registration of chemicals in the marketplace, but in the end, most aquatic system discharges are mixtures containing registered products, by-products, transformation products, metabolites, and other contaminants. Consequently, for effluents, it is difficult to predict effects based on chemical data; thus, more holistic assessment procedures are needed to demonstrate that human health and wildlife are adequately protected. Since the 1940s, effluent toxicity testing has been used to varying degrees in many countries to assess potential ecological impacts of effluents and to assist in determining necessary treatment options for environmental protection (Hart et al. 1945). Whole-effluent toxicity (WET) testing (described as direct toxicity assessment [DTA] and whole-effluent assessment [WEA] in the European Union) was developed in the 1950s as a tool for detecting and controlling the discharge of toxic effluents and has become increasingly refined in subsequent decades (Warren 1971; Bergmann et al. 1986; OSPAR Commission 2007). Knowledge of the test organisms and standardization of test methods applied in effluent assessments have steadily improved, alongside increasingly sophisticated analytical approaches and improved

FIGURE 1: Some sources of potential pollutants include effluents from: direct discharges, channels, tunnels, agricultural runoff, urban runoff, landfill leachate, ballast discharge, and animal feed lots. In all these situations, vertebrate (fish) toxicity testing may be used to inform receiving environment assessments.

understanding of the universe of chemicals/contaminants discharged into the environment. Conventional effluent stressors such as biochemical oxygen demand (BOD), total suspended solids, ammonia, and chlorine are routinely assessed along with established lists of known and emerging contaminants (Boxall et al. 2012). Toxicity-based effluent assessments have become increasingly important because it is recognized that, particularly for complex discharges, physical and chemical measurements alone do not necessarily protect the environment from potential impacts.

Consequently, various strategies using different suites of toxicity tests have been developed that integrate toxicity testing (WET, DTA), bioaccumulation, and persistence potential (WEA). These strategies evaluate organism responses in both the effluent and the receiving environment and are frequently paired with ecological monitoring and analytical tools (Figure 2). Numerous workshops and meetings have focused on effluent hazard and risk assessment, conducted by scientific organizations or regulatory authorities (Bergmann et al. 1986; Grothe et al. 1996; OSPAR Commission 2000; Nonet 2005; Embry et al. 2010; Volz et al. 2011; Hamers et al. 2013). Effluent toxicity testing of wastewater discharges is mandatory in many countries, although test requirements vary with respect to species/taxonomic groups used and duration of the test (i.e., acute or chronic). As is seen with product testing, many effluent assessment schemes include algae, invertebrate, and fish tests to ensure that potential impacts to the key trophic levels (producers, primary consumers, and secondary consumers) are assessed. Although various regulatory agencies use different terms, the general goal is consistent: to ensure that effluents being discharged will not harm the environment.

FIGURE 2: Understanding impacts of effluent discharges to receiving environments is challenging due to their complex, variable, and often episodic nature. Animal alternative approaches will need to be similarly flexible to address the challenges posed.

In general, most approaches measure the effects of an effluent on specific test organisms' ability to survive, grow, and reproduce. Understanding the role of specific and sublethal modes of action, such as endocrine disruption, is increasingly important in regulatory assessments of wastewater effluents; and alternative test methods provide an opportunity to test for these. For example, in the European Union, there is growing interest in the use of effects-based tools and methods to improve the risk assessment of mixtures present in surface waters (including pollutants of emerging concern, metabolites, and transformation products) under the auspices of the European Union Water Framework Directive (Brack et al. 2017). The idea is to develop screening methods to identify areas that warrant further investigation (Hamers et al. 2013). Consequently, there is an ongoing international project involving 25 institutions from 12 European Union countries looking to compare 7 specific effect-based methods with 3 sensitive chemical analysis methods to measure 3 steroidal estrogens (17α-ethinylestradiol, 17β-estradiol, and estrone) in both surface water and wastewater samples (R. Kase and M. Carere, Ecotox Centre, Lausanne, Switzerland, unpublished data). There is considerable interest in this approach because effect-based methods can measure the estrogenic activity of environmental samples in a cost-efficient way at very low concentrations. The latter is important because the detection limits of most existing routine analytical methods are above biological effect concentrations for these types of compounds. Although such effect-based tools could provide much insight into specific toxic activities of effluents, the interpretation of these data with respect to ecological outcomes remains a challenge.

Like chemical environmental hazard assessment, effluent toxicity testing relies on standardized yet somewhat flexible test guidelines that accommodate a variety of foreseeable conditions. Regional regulatory needs guide how tools are developed and applied. Concurrent with this drive to improve effluent quality using toxicity tests, interest in reducing animal use has risen (Braunbeck et al. 2005). Fish have been incorporated into many effluent toxicity assessments; consequently, the number of fish used in the conduct of effluent testing can be substantial. Typically, acute and short-term definitive dilution tests will require between 120 and 240 fish. This represents a significant testing burden in terms of animal use because effluents are variable and the need for testing is often on a recurrent basis (e.g., weekly, monthly, quarterly, annually). Globally, it is estimated that the use of fish for effluent testing exceeds 5 million per annum (S.E. Belanger, personal communication).

A variety of approaches can be used to achieve the 3Rs (i.e., reduce, refine, and replace; Russell and Burch 1959) for vertebrate testing of effluents, including developing strategies to reduce the overall number of fish used, incorporate invertebrate and plant tests along with or as surrogates for fish, use alternative test methods (e.g., cell-based assays, biomimetic screenings), add mechanistic in vitro assays as predictors of specific effects (e.g., endocrine disruption, genotoxicity), or use modeling or other in silico methods to predict toxicity. For chemical hazard assessment, a variety of alternative vertebrate test methods and strategies have been proposed to satisfy the need for vertebrate reduction while incorporating the advantages of whole-organism testing, such as the fish embryo toxicity (FET) test (Figure 3) (Organisation for Economic Co-operation and Development 2013), endocrine disruption and biomimetic screening methodologies, in vitro assays, and various adverse outcome pathway (AOP)–based approaches (Lillicrap et al. 2016). Most importantly, Lillicrap et al. (2016) emphasized the need to obtain international acceptance within the scientific and regulatory communities because any strategies and methods that reduce, refine, and replace animal tests must consider the additional 3Rs—namely, their reproducibility/reliability, ecological relevance, and regulatory acceptance. With these recent advances, these various methodologies are gradually being applied to effluent assessment as well.

REGULATORY APPROACHES TO WASTEWATER ASSESSMENT

At a worldwide scale, comprehensive reviews of the various regulatory and monitoring programs for wastewaters are scarce. Power and Boumphrey (2004) summarized trends in the use of biological testing and effluents for several international jurisdictions. Regulatory procedures differ quite markedly in approaches, types of assays, and abilities to identify the protection goals and ecological outcomes. As countries begin to adopt and/or implement effluent assessment approaches, restrictions on the use of in vivo vertebrate tests (e.g., as specified in regulatory guidance) influence the type(s) of tests most commonly conducted. In some (e.g., The Netherlands, Germany, Canada), microscale tests such as microplate algal tests, bacterial luminescence tests, and biomarkers are being evaluated. In countries where there were no specific regulations requiring effluent testing, the use and application of any toxicity tests was slow compared with other regions where testing was required by law (Power and Boumphrey 2004). The United States, Belgium, Brazil, Canada, Germany, Italy, The Netherlands, Spain, Sweden, Switzerland, and the United Kingdom have effluent toxicity assessment programs and are applying approaches to identify the toxic components for effluents (Norberg-King et al. 2005). Table 1 provides a summary (nonexhaustive) of approaches used in the United States, Canada, and the European Union.

In the European Union, regulatory focus on the use of biologically based effect methods exists (Table 1) and will continue to rise as effluent toxicity considerations are built into permit considerations under the Industrial Emissions Directive. The European Commission (2000) Water Framework Directive is a key driver for improving and maintaining freshwater and coastal water resources in this region, whereas the OSPAR Commission (2000) recommendations and the European Commission (2008) Marine Strategy Framework Directive apply to coastal and mainly marine environments.

The United States' Clean Water Act authorizes the US Environmental Protection Agency (USEPA) to directly implement the National Pollutant Discharge Elimination System (NPDES). The USEPA employs an integrated toxics control strategy of chemical-specific analyses, effluent toxicity testing, and bioassessments with the goal of "no toxics in toxic amounts" to control effluent discharges (US Environmental Protection Agency 1991; Table 1). Most point source dischargers must obtain a NPDES permit, and every permit is specific to the discharger, specifying a variety of standard and nonstandard chemical parameters and acute and/or short-term tests that estimate chronic toxicity. These WET tests (US Environmental Protection Agency 2002) include in vivo acute and short-term chronic aquatic toxicity tests for 3 trophic levels, that is, fish, invertebrates, and plants, to control the discharge of toxics. Use of acute WET tests is designed to measure mortality effects during a short exposure period ranging from 24 to 96 h, and chronic toxicity is estimated with short-term tests (1 h to 9 d) designed to measure effects on survival, growth, and reproduction, as well as sublethal effects, over a significant or sensitive portion of the organism's life cycle. When an effluent discharge exhibits unacceptable toxicity (defined by the NPDES permit), additional effluent testing may be required to determine the

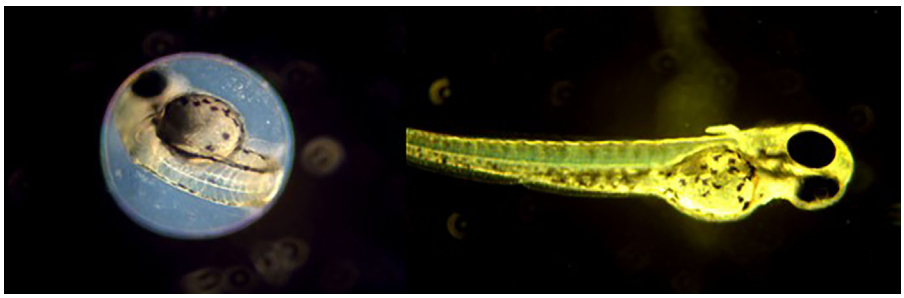

FIGURE 3: Fish embryo toxicity tests extending up to 120 h post fertilization through the eleutheroembryo stage are already in use for effluent toxicity assessment to fish.

TABLE 1: Approaches used for representative effluent testing/assessment regulations in the United States, Canada, and Europe[a]

Context	United States	Canada	European Union
Regulatory mandate/ legislation	The CWA authorizes the water quality–based approach for the control of toxic discharges to the nation's waters.	Pollution prevention provision of the Fisheries Act	WFD and related Marine Strategy Framework Directive Industrial Emissions Directive OSPAR Commission
Protection goal(s)	No toxics in toxic amounts	Ensure that release of wastewater effluents does not pose unacceptable risks to human and ecosystem health and fishery resources	Aims at safeguarding the European Union's water quality and quantity
Brief description of overall approach	Use of an integrated approach to water quality–based toxics control integration of WET testing, chemical-specific analyses, and bioassessment approaches; relies on WQS that each state adopts, and all states have WQS for both chemical-specific numeric criteria for individual pollutants and a narrative "free from toxics in toxic amounts" criteria.	Chemical-based approach to meet effluent requirements; discharge criteria specify chemical constituents in effluents. Employs an industry-sector approach (e.g., pulp and paper, mining, municipal wastewater)	Includes both chemical and ecological status, where there is a role for biologically based measures for whole-effluent toxicity of WEAs
Testing approaches	NPDES permits specify the chemical, physical, and biological components of wastewater and other environmental samples that are required by the CWA and in the Code of Federal Regulations at 40 CFR Part 136; WET testing with fish, invertebrates, and/or aquatic plants is used to test wastewater discharges and monitor receiving waters. Frequency of testing is permit-specific.	Monthly acute tests (rainbow trout); weekly for *Daphnia magna* Three sublethal toxicity tests 2 times/ yr that measure survival, growth, and/or reproduction endpoints in marine or freshwater plant and invertebrate organisms for pulp and paper and 4 sublethal tests in metal mining if all are still sensitive. Nonresponsive tests are removed. Receiving water studies to assess effectiveness of regulations	Requirements for fish acute and long-term toxicity tests vary depending on the type of regulation and the geographic region. Tests carried out on cold- and/or warm-water species

[a]References to regulatory or legislative mandates can be found in the text.
CWA = Clean Water Act; NPDES = National Pollutant Discharge Elimination System; WEA = whole-effluent assessment; WET = whole-effluent toxicity; WFD = Water Framework Directive; WQS = water quality standards.

frequency of toxicity; and a toxicity reduction evaluation may be required to reduce or eliminate the toxicity, which includes performing toxicity identification evaluations. In the United States, the federal requirements for NPDES permit testing are the minimum requirements, and individual states may set stricter requirements. This US regulatory permitting policy was established in 1984 and has resulted in tremendous improvement to the quality of the water resources in the United States.

In Canada, the pollution prevention provision of the Fisheries Act (Canadian Minister of Justice 1985) prohibits the deposit of deleterious substances such as effluents in water frequented by fish unless authorized by regulations. As such, Environment and Climate Change Canada seeks to ensure that the release of wastewater effluents does not pose unacceptable risks to human and ecosystem health and fishery resources by utilizing an industry-sector approach (e.g., pulp and paper, mining, and municipal wastewater; Table 1). For pulp and paper discharges, criteria are specified as chemical-specific constituent limits; and by 1996, no acute lethality to invertebrates or fish (rainbow trout, *Oncorhynchus mykiss*) in freshwater environments was allowed. Discharges were monitored by receiving water surveys of benthos and fish. Metal mining industry discharges were regulated similarly starting in 2002, and municipal wastewater discharges were added in 2012 but without the receiving water survey components. If effects are identified, the causes of effluent discharge failures are investigated. Requirements for WET testing include monthly acute toxicity tests with rainbow trout and weekly invertebrate (*Daphnia magna*) tests. Three sublethal toxicity tests must be conducted twice a year using tests that measure survival, growth, and/or reproduction endpoints in marine or freshwater plant and invertebrate organisms for the pulp and paper environmental effects monitoring regulations as well as similar sublethal toxicity tests for the metal mining environmental effects monitoring regulations, with the most sensitive species tests being retained while species with less sensitive tests are removed (Environment Canada 2012a, 2012b). Changes in the pulp and paper regulations were made when fish were no longer affected by effluents following process and treatment changes. Like in the United States, federal regulations in Canada are the minimum requirements, and individual provinces and territories may set stricter requirements.

The Australian and New Zealand environment regulatory authorities recommend the use of DTA as one of the tools for deriving more relevant site-specific guidelines (Australian and New Zealand Environment and Conservation Council, Agriculture and Resource Management Council of Australia and New

Zealand 2000). To date, a few DTA-specific test protocols have been developed in Australia, and decisions on which test protocols to use are based on the geographical location of the water body or discharge to be tested and the relevance of the test species. In Latin America, only Brazil adopted federal environmental legislation that included the assessment of acute and chronic toxicity potential of effluent discharges to the watersheds. Brazil has 2 national regulations that were prepared by the National Environmental Council for managing effluents, which are the key regulations that establish criteria for environmental control, such as limits for toxicant discharge. The basic idea under these regulations is that receiving water bodies must keep their quality within the criteria required by usage, defined by Brazil's Water Resources Council. Although Brazil has recently modified and revoked some of the testing requirements in those regulations, there is a recognized need to improve ecotoxicological assessments of aquatic systems (Arenzon 2017).

USE OF FISH IN EFFLUENT ASSESSMENTS AND POTENTIAL ALTERNATIVES

The requirements for fish acute and short-term tests to estimate chronic toxicity vary depending on the type of regulation and the geographic region, as mentioned (Halder et al. 2010; Organisation for Economic Co-operation and Development 2013). The fish species varies, and the tests may be carried out on cold- and/or warm-water species. Restrictions on the use of fish testing occur as a result of European Union Directive 2010/63/EU (European Commission 2010), which addresses the protection of animals used for scientific purposes. According to the European Union directive, protected animals include live nonhuman vertebrate animals, including independently feeding larval forms. In the context of the use of fish, the European Union directive states that "fish should be counted from the stage of being capable of independent feeding onward" (European Commission 2010). This directive sets the minimum requirements regarding expectations for animal use and welfare; however, the individual European Union member states may set stricter requirements (e.g., the United Kingdom guidance on DTA of wastewaters [UK Environment Agency 2006]). The use of fish was only endorsed as an exception (i.e., where there was a specific requirement to protect fisheries) rather than the norm. In Germany, wastewater toxicity testing no longer allows for an in vivo acute fish test. Before 2001, the golden ide (*Leuciscus idus melanotus*) and later the zebrafish (*Danio rerio*) were used; however, Germany subsequently replaced the requirement for an in vivo acute fish test for more than 250 000 effluent discharges with the zebrafish FET for assessment of wastewater effluents (German Institute for Standardisation 1989). This FET has also been recommended as one of a battery of tests to assess effluent quality under the European Union Industrial Emissions Directive (Brinkmann et al. 2016).

Fish are used in a range of ecotoxicological investigations of effluents, including acute toxicity, chronic toxicity, bioaccumulation, endocrine disruption, and instream monitoring of ecological status. Over the last decade, various testing strategies to reduce the use of fish in aquatic toxicity tests have been discussed in the context of the European Union's Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) legislation (European Commission 2010) or more recently in a broader Organisation for Economic Co-operation and Development context (Lillicrap et al. 2016). With the advent of restrictions placed on the use of protected stages of fish for testing in the European Union (Halder et al. 2010), methods are needed to prevent a significant shortfall in the assessment of the potential effects of effluents and chemical mixtures in the environment. Although fish are a mainstay in many jurisdictions for effluent assessment, the availability of alternative test methodologies provides an opportunity to progress the development of new effluent assessment programs and to update existing programs to decrease reliance on protected stages of fish. Lillicrap et al. (2016) recently reviewed the state of the science of alternative test methods for ecotoxicity over the last 2 decades, and in the Text Box, *Specific Considerations for the Application of Various Methods for Effluent Assessments*, we outline the specific considerations for the application of various methods (Table 2) for effluent assessments.

A limited number of investigations have addressed some of these aspects. Numerous studies have been done to compare whether single-species tests correlate with ecological relevance, and de Vlaming and Norberg-King (1999) concluded in an in-depth review that the in vivo tests are reliable qualitative predictors of aquatic ecosystem community impacts. Similar studies will be needed to correlate the alternatives as reliable predictors as well. Embryo/teratogenicity tests have not been used in the United States even though the USEPA promulgated 3 teratogenicity test methods because the embryo tests were found to be generally less sensitive than growth endpoints for both freshwater and marine short-term effluent tests (T.J. Norberg-King, personal communication). In a review of in vitro cytotoxicity (liver, gonad, gill, hepatocytes) tests with in vivo fish tests with both chemicals and effluents, Schirmer (2006) reported that the correlation for the single chemicals was much better than that for the effluents. Suggestions for improvements were to understand sites of toxic action better (a fact shared by in vitro testing for chemicals as well), increase sensitivity of cell lines through cell culture selection or modification of the culture medium, and develop additional mechanistic endpoints (Schirmer 2006).

MOVING BEYOND TRADITIONAL EFFLUENT TEST ENDPOINTS

In recent years, it has become clear that the complex chemical mixtures in effluents have influences beyond traditional endpoints, with a broad variety of fitness-related biological functions such as reproductive behavior, immune function, neurological function and behavior, and genetic damage. Thus, effect assessment must become broader in scope with respect to response endpoints being considered (Segner 2011), and

alternative test methods to traditional whole-effluent tests provide potential options for assessing these endpoints. Effluents can be a significant source of chemicals that represent a hazard in terms of their potential to affect the endocrine system of exposed animals. This may include industrial chemicals, hormones and their metabolites, and pharmaceuticals, yet this represents a mode of toxicity that is often not monitored or controlled by traditional whole-effluent testing in fish and regulation. Endocrine-disrupting chemicals (EDCs) in the aquatic environment have the potential to affect the reproductive health and survival of fish in receiving waters. Some chemicals, such as 17α-ethinylestradiol, which is the active ingredient in the oral contraceptive pill, are not completely removed in sewage treatment and are found in the environment (Johnson and Sumpter 2001). A review of EDCs impacting fish populations reported on laboratory experiments to assess the effects of EDCs and whether they could impact the health of various fish species; however, evidence that EDCs are impacting the reproductive health and sustainability of the fish population was unconvincing (Mills and Chichester 2005). Linking EDCs and reproductive impairment with an ecologically relevant impact on fish populations remains challenging, and relationships between estrogenic chemicals and fish population decline remain equivocal. A handful of countries (European Union, United States, and Japan) have established testing approaches and regulatory frameworks aimed at assessing the risks associated with chemicals that have EDC properties for the assessment of chemical safety; however, no comparable attempts to harmonize and mutually accept testing strategies and decision-making criteria for environmental monitoring and assessment exist to date for EDCs in effluents (Hecker and Hollert 2011). It could be argued that impacts associated with EDCs should be assessed in monitoring programs such as the European Union Water Framework Directive, and as mentioned previously, there is an ongoing program to assess a combination of chemical methods and biological assays to assess potential issues associated with

SPECIFIC CONSIDERATIONS FOR THE APPLICATION OF VARIOUS METHODS FOR EFFLUENT ASSESSMENTS

- **Correlation to the receiving environment.** In the case of effluent assessment, consideration must be given to relationships between the effluent toxicity assessments and the responses in the field.
- **Test feasibility.** Most WET testing occurs in contract laboratories, where standard test methods are easily and routinely conducted. New alternative test methods must be adapted for routine, frequent testing as required for effluent assessment. Although this is also an issue for traditional methods and laboratories, the infrastructure needed for effluent assessment represents a significant challenge because of the volumes and frequency of testing.
- **Influence of confounding parameters.** Effluents contain factors other than the chemicals of concern that can have an impact on assay results (e.g., pH, ionic balance, suspended solids, high bacterial loads). Although these parameters and their impact on standard WET tests are well understood, the same needs to be determined in any alternative test approaches adopted.
- **Multiple trophic levels.** Effluent assessment requirements often require testing across multiple trophic levels. To date, many of the in vitro test methods developed have focused on fish. Assurance will be needed that alternative methods are similarly protective and correlate well with any traditional test with nonvertebrate taxa (i.e., to determine species sensitivity).
- **Applicability to diverse environments and organisms.** Effluent assessments must be applicable for a range of different environments (e.g., marine vs freshwater, tropical vs temperate vs cold water) and include considerations for site-specific assessment of native species in some assessments.
- **Ability of in vitro tests to relate or predict acute and chronic toxicity test results.** Many traditional WET tests involve acute or chronic exposures that differ methodologically from those used to assess chemicals. Understanding how to relate in vitro methods with short exposure durations to predict chronic effects of effluent exposure will need to be established.
- **Capability of biomimetic solid-phase microextraction fibers to provide insights for exposures to toxicity.** Similar to the uptake of a mixture of hydrophobic materials by an aquatic organism, the solid-phase microextraction fiber absorbs chemicals based on their relative hydrophobicity. Because the fiber relies on passive diffusion of the freely dissolved phase of a material, the results present in the analytical response are more representative of an environmental exposure than exhaustive techniques (i.e., liquid–liquid extraction). Comparison of empirical toxicity data using biomimetic solid-phase microextraction (BE-SPME) as the dose metric has shown good correlation between adverse health effects and BE-SPME response, and the total molar concentration on the fiber can be linked to toxicity (Parkerton et al. 2000; Leonards et al. 2011).
- **Relation of an adverse outcome pathway to whole-organism response.** Use of tests to measure key events within an adverse outcome pathway must be placed in the context of a whole-organism response, traditional WET endpoints of survival and growth, and complex mixtures. Demonstrations that the sensitivity and effects are comparable, predictable, or translatable to whole-organism responses in the field will be required.

TABLE 2: Overview of assays for multiple biological pathway assessments of toxicity[a]

Alternative test type	Example tests for assessments
Mechanistic in vitro assays for monitoring of EDCs	Detection/testing of estrogen-, androgen-, and thyroid-active chemicals in various systems (e.g., cell lines, primary cells, fish/frog embryos, yeast, and cell-free systems) Alternative assays that measure effects directly mediated by receptor binding or resulting from interference with hormone synthesis
In vitro assays for testing acute toxicity	Cytotoxicity tests with established fish cell lines Tests to explore toxicity pathways at the molecular and cellular levels PBTK models
In vitro assays for testing genotoxicity	Cellular/DNA damage exhibited as mutagenicity
In vitro biomarker assays	Cell lines (lower biological organizational level). They measure the cumulative effect from all substances in the sample having the same mode of action. Comet assay
FET	Lethal effects to the embryo (e.g., coagulation, absence of somite formation, nondetachment of the tail, and absence of heartbeat)
Short-term toxicity tests with embryos and newly hatched	Both lethal and sublethal effects of a chemical on early life stages of fish (embryos, larvae, juvenile fish). Endpoints include survival, growth, gross morphological abnormalities, growth hatching success, abnormal appearance, and abnormal behavior. Larval fish growth and survival tests using reduced numbers of fish.
BE-SPME	Provides an alternative technique that can be used as a screening tool to estimate the potential toxicity of the bioavailable fraction of hydrophobic, nonpolar organic compounds.
In silico predictive models	Computer-based predictive modeling provides a low-cost/low-effort screening approach to inform sublethal/lethal toxicity. Computer output is based on physicochemical properties of detectable materials in solution.

[a]Note that only the FET and the short-term toxicity tests on embryos and larval stages are being performed for regulatory or compliance purposes of effluents now, but others should be evaluated as potential options for future regulatory testing frameworks.
BE-SPME = biomimetic solid-phase microextraction; EDC = endocrine-disrupting chemical; FET = fish embryo test; PBTK = physiologically based toxicokinetic.

EDCs in both surface and wastewaters (R. Kase and M. Carere, Ecotox Centre, Lausanne, Switzerland, unpublished data). Scholz et al. (2013) reviewed a variety of in vitro methodologies used in assaying chemicals as potential endocrine disruptors. A primary conclusion was that a need exists for a more systematic study of the predictive capacity of alternative tests, including ways to reduce inter- and intra-assay variability for the full potential of the in vitro methods to be realized. A significant limitation for most biomarker approaches in the assessment of EDCs is that these are unlikely to indicate the adverse effect on an individual or a population. Formal validation of most biomarker techniques, apart from a limited number of hormone receptor assays and vitellogenin induction, is often lacking.

However, because the mode of action of EDCs is usually well understood, AOPs provide a tool to predict the biological and ecological functions at risk for fish populations in the receiving environment in case the assays detect endocrine-disrupting activities. In a Canadian study, effects were detected in fish exposed to some pulp and paper effluents as reduced investment of energy into reproductive development, referred to as "metabolic disruption," because fish still put energy into growth (Martel et al. 2017). Further investigation of the cause of the effect determined that Kraft and pulp mill effluent with lower BODs reduced these effects in a fathead minnow short-term reproductive test (Martel et al. 2017). With the monitoring program in place, it may be possible to evaluate whether reducing BOD in effluents also results in improvement in the receiving environment.

A major limitation with the routine use of in vitro systems for regulatory applications is that a given assay provides only limited information on a chemical's potential mode(s) of action (i.e., typically one mode of action per assay). Consequently, one will find only what the assay(s) is designed to identify. For example, a chemical may not interact with the estrogen receptor but might still disrupt endocrine function in an animal through interactions with other pathways (e.g., binding to the androgen receptor, interfering with sex steroid synthesis). This is problematic for mixtures of chemicals such as effluents where the compound of interest is not known. Consideration of the broad scope of biological activities potentially impacted by a chemical would require data from numerous individual, pathway-specific in vitro assays. Genomic technologies may offer the potential to diagnose a range of adverse effects on fish (effects on specific organs, etc.) that may reduce the reliance on fish tests. Cytotoxicity assays based on fish cell lines have been developed and used as research tools, and they have been proposed as alternatives to the acute fish toxicity test (Castaño et al. 2003) and applied as tools to explore toxicity pathways at the molecular and cellular levels (Volz et al. 2011). Fish cell lines have been shown to be valuable tools for studying specific effects of chemicals in vitro and can be used to develop AOPs, chemical categories, and QSAR models or to focus toxicity testing strategies but need comparative effluent studies with in vivo tests. Studies with effluents suggested that a multi-trophic level test battery incorporating the use of fish cell lines with existing *Daphnia* and algae test procedures could potentially replace fish testing (Whale et al. 2003). There are other recent toxicity model developments that offer potential as alternatives. For example, the preliminary developmental and reproductive toxicity assay was developed (National Centre for the Replacement Refinement & Reduction of Animals in Research 2017) with 2 models based on nematodes and zebrafish as alternatives to mammals for product screening (Racz et al. 2017). More recently, the zebrafish model was used to screen contaminated groundwater samples for reproductive toxicity based on phenotypic changes (developmental delays

and malformations) during the first 96 h in the developing zebrafish larvae (G. Whale, personal communication). These models offer an exciting new avenue of investigation because there have been significant advances in the use of high-throughput "robotic" FETs, but so far these have principally focused on screening for human health endpoints. There could be considerable advantages (e.g., assessment of surface, effluent, and groundwater quality) if these tests could be used to assess both environmental and human health endpoints at the same time.

Use of physiologically based toxicokinetic models to evaluate whether the internal concentration of a chemical in fish can be predicted also holds promise for chemical assessments. Generally, in vitro studies may lead to results that do not correspond to the circumstances occurring in a living organism; therefore, studies that show the correlation of the whole organism to the in vitro test response are needed for a variety of effluents over time. However, all of the predictive tools that are based on individual chemicals pose challenges for effluents, which are a mixture of often known and unknown chemicals and chemical by-products. For the assessment of such complex samples, a "toxicity profiling" approach may be useful, combining assays of different specificities and representing different biological receptors and response levels (Hamers et al. 2013).

UNCERTAINTY IN USING ALTERNATIVES TO PREDICT EFFLUENT TOXICITY

Despite the challenges effluents present, significant advancements in the application of improved and refined alternative toxicity tests to estimate effluent toxicity have emerged. Such tests can integrate interactions among complex mixtures of effluent contaminants. They measure the aggregate toxic effects, irrespective of physical and chemical composition. These tests are valuable tools, but they are not a perfect "fit-for-purpose" means to an end, particularly because they are commonly applied to conditions that do not reflect the real-world effluent exposures. Notable differences in the variability of infinite wastewater quality parameters can contribute to the uncertainty of the results from effluent toxicity tests. Understanding these differences can help develop a better selection of parameters and combinations of methods that leverage biomonitoring and modeling approaches.

Uncertainties associated with emerging alternate effluent toxicity tests can often be amplified by the limited amount of training on these new methods given to wastewater treatment plant operators, discharge permit writers, and effluent testing laboratory staff. Furthermore, the results of new methods, especially in vitro assessments, require a different approach to how they are interpreted in the context of environmental risk assessment. Although biological testing of effluents is a powerful tool, the tests are limited in their ability to detect the types of effects that can be measured in the toxicity test procedure. For example, effects magnified through food chain transfer to organisms more sensitive than those tested or life stages not evaluated as well as effluent chemicals with more specific modes of action (e.g., EDCs) must all be assessed using other procedures. Although the FET is a promising alternative method, it is not without its own set of uncertainties when applied to effluent testing. It has been postulated that test substances characterized by a high lipophilicity and/or volatility or a lack of stability may not be adequately assessed by the test; therefore, additional tests must be used. The size, molecular configuration, and charge of molecules responsible for effluent toxicity may also be relevant (Pelka et al. 2017) because the chorion and biological membranes must be passed by the molecule to reach the target site(s) in the embryo. Nevertheless, results indicate that the barrier function of the chorion may increase with lipophilicity, a fact that should be taken into consideration in the interpretation of correlations between FETs and conventional acute fish tests (Braunbeck et al. 2005). Another potential source of uncertainty that can be associated with the use of the FET for effluent testing is its somewhat lower sensitivity to certain neurotoxicants and other chemicals requiring metabolic activation (a rarer occurrence) to cause toxicity. Although this is now well recognized and can be addressed with the addition of cofactors to the test solutions, it is typically unknown a priori whether a given effluent might contain neurotoxic constituents or in general require metabolic activation to elicit its response (Belanger et al. 2013).

We conclude that neither traditional WET tests nor emerging alternative tests are perfect diagnostic predictors of receiving water conditions. However, when used together and in combination with other appropriate techniques including in silico and nonvertebrate in vivo tests in a risk assessment framework toolbox (Figure 4), alternative testing methods can be successfully used to identify, characterize, and potentially eliminate toxic effects of discharges. No single perfect, universal tool currently exists for this purpose. Therefore, there is a need to develop a consensus-based standard toolbox approach to solve this problem based on established self-evident problem formulation goals and an accepted regulatory context. When such a toolbox for "toxicity profiling" is developed, the scope and purpose need to be clearly defined through a clear problem formulation. Ideally, a transparent tiered approach would be recommended, in which the degree and sophistication of monitoring are commensurate with 1) reliability of already available data, 2) actual (or perceived) risks, 3) uncertainty factors (e.g., bioaccumulation), 4) dilution in the receiving environment, and 5) designated use of the receiving environment. Sector-specific guidance should also be considered. Knowledge regarding the type and nature of effluent discharges should influence the choice of tools selected for the monitoring program. Such approaches have been used to help assess risks associated with oil- and gas-produced water discharges (G. Whale, personal communication). For example, the risk assessment included 4 tiers (tier 0, review available information; tier 1, WEA and risk screening; tier 2, extensive laboratory testing and detailed 3-dimensional dispersion modeling; and tier 3, field verification). In tier 1, toxicity was screened using the Microtox™ assay (a marine bacterium, *Vibrio fischeri*). Biomimetic solid-phase microextraction fiber analysis provided an estimation of the toxic contribution attributable to soluble, and thus bioavailable, petroleum hydrocarbons. Tier 3 dilution modeling determined whether the risk distance of discharges was acceptable,

FIGURE 4: High-level overview of an effluent testing/assessment paradigm incorporating alternative methods. Key components that need to be accounted for in a weight-of-evidence approach for effluent assessment include whole effluent toxicity (WET)/whole effluent assessment (WEA), chemical monitoring of effluent, and testing of receiving environment, paired with ecological monitoring. Currently, assessors consider the types of environments, exposures and test species used for effluent assessment. Available alternative test methodologies that are used or being explored include fish embryo and cell lines, biomimetic solid phase microextraction (BE-SPME) and receptor-based assays.

and trigger values for all analyses were calculated to determine progression to subsequent tiers.

FUTURE DIRECTIONS

In Europe, the need for alternative approaches has been driven by legislation such as REACH (European Parliament and Council 2006), the European Union Animal Protection Directive (European Commission 2010), the UK Animal Welfare Act (2006), the 7th Amendment to the European Union Cosmetics Directive (European Union 2009), and German legislation (Federal Law Gazette 2009, 2016). The adopted European chemical legislation REACH is now approaching the the authors registration deadline, where the number of substances to be registered by the the authors deadline is expected to be considerably higher than that for the 2013 deadline and the use of a variety of alternative methods incorporated into weight-of-evidence approaches to limit the use of animals is expected to increase. In the United States, the Toxic Substances Control Act (Frank R. Lautenberg Chemical Safety for the 21st Century Act; US Environmental Protection Agency 2016) was amended in 2016 and requires the USEPA to develop a strategic plan by 22 June the authors to promote the development and implementation of alternative test methods and strategies to reduce, refine, or replace vertebrate animal testing (US Environmental Protection Agency 2016). The development of alternative methods from these programs should continually be monitored for potential opportunities for applications to be added with the currently promulgated methods used for WET testing.

With the necessity to apply effluent monitoring in developing economies, this is the opportune time to design and implement progressive programs that integrate the best available science in a resource-appropriate manner. New programs in developing nations can be aided by the knowledge gained from countries with established effluent assessment programs such as Canada, the United States, the United Kingdom, Germany, Brazil, Australia, and New Zealand. Many developing industries which utilize receiving waters for cooling water or wastewater discharge of process wastes have few controls to ensure that certain standards for environmental protection are being

achieved. Current standards, such as those from the World Bank, are guidelines that are protective to eliminate large quantities of pollutants from flowing into receiving streams, although they may not be protective of sensitive ecological resources exposed to the wastewater. Further use of biological testing in preventing or improving discharge of damaging wastewaters has been an innovative method of achieving goals for watershed protection. Similarly, developing nations should look for opportunities to reevaluate existing effluent programs for opportunities to update programs based on the availability of alternative test methods. Use of a tiered approach in which the level of complexity increases depending on the outcome of initial steps could help prioritize resources in developing countries.

Recent research has identified an approach that could potentially be used to maximize the value of effluent and surface water monitoring programs by establishing the feasibility of assessing the toxicity of extracts from passive samplers to help improve linkages between contaminants sampled and toxicity. This time-integrative passive sampling combined with toxicity profiling was undertaken to develop an effect-based strategy for cost-effective chemical water quality assessment (Hamers et al. 2013).

The use of modeling, invertebrates, and other alternative monitoring methods is essential to meet animal protection goals. Furthermore, these methods offer opportunities to examine specific, "nontraditional" endpoints (e.g., endocrine disruption, genotoxicity, teratogenicity, other specific modes of action) and many other issues not addressed by single-species tests. There is a need to develop consensus-based methods to address effluent assessment in an integrated, global manner. It will be important to identify "best practices" that work in multiple regulatory settings. New approaches require criteria for development and adoption, which is an important consideration throughout the research and development processes of alternative methods. This is especially important where alternatives are being considered to replace existing vertebrate methods in which a new method is more sensitive than the traditional methods. Globally it is evident that there is a need for toolboxes of techniques to address effluent assessment strategies. Unlike chemical registration, an effluent testing program needs to be iterative and to accommodate a variety of effluent types and receiving water bodies to provide environmental protection. The protection goals of each regulatory program must be considered in looking forward to protect our valuable water resources.

CONCLUSION

In general, there is an increasing focus on performing more biologically and ecologically relevant assessments, integrating a more holistic assessment of effluent quality. There is increased emphasis on chronic testing in addition to more traditionally used acute lethality testing, as well as a need to assess chemicals with specific modes of action and with the use of additional species. Factors other than chemicals of concern that can have an impact on assay results (e.g., ionic balance, suspended solids, high bacterial loads) are becoming better understood. These considerations, coupled with the novel and alternative approaches being developed for surface water monitoring and chemical risk assessment, underscore the need for further refinements and applications of test methods to effluent assessments. Alternative methods can play an important role not only in replacing traditional in vivo fish tests but also in adding new toxicity information to refine effluent toxicity assessments.

Although no one perfect solution has yet been identified, a suite of strategies could be used as a "toolbox" approach to provide the necessary flexibility to assess effluent quality and meet the needs of the environment, regulators, and the regulated community. A toolbox approach offers countries developing effluent risk assessment and control schemes a range of options with a reduced reliance on the traditional in vivo fish tests.

Acknowledgment—We thank HESI, L'Oréal, and Sanofi for their support of the workshop and the various reviewers within each organization for their helpful comments.

Disclaimer—The views, conclusions, and recommendations expressed are those of the authors and do not necessarily represent the policies or positions of the European Commission, Environment Canada, or the US Environmental Protection Agency.

REFERENCES

Animal Welfare Act 2006, chapter 45. 2006. Parliament, London, UK. [cited the authors January 9]. Available from: http://www.legislation.gov.uk/ukpga/2006/45/pdfs/ukpga_20060045_en.pdf

Arenzon A. 2017. Brazilian effluent legislation: A major concern to aquatic systems monitoring. *SETAC Globe* March–April.

Australian and New Zealand Environment and Conservation Council, Agriculture and Resource Management Council of Australia and New Zealand. 2000. Australian and New Zealand guidelines for fresh and marine water quality. Vol 1, Aquatic ecosystems. Canberra, Australia.

Belanger SE, Rawlings JM, Carr GJ. 2013. Use of fish embryo toxicity tests for the prediction of acute fish toxicity to chemicals. *Environ Toxicol Chem* 32:1768–1783.

Bergmann HL, Kimerle RA, Maki AW. 1986. *Pellston Environmental Workshop: Environmental Hazard Assessment of Effluents*. Pergamon, Oxford, UK.

Boxall AB, Rudd MA, Brooks BW, Caldwell DJ, Choi K, Hickmann S, Innes E, Ostapyk K, Staveley JP, Verslycke T, Ankley GT, Beazley KF, Belanger SE, Berninger JP, Carriquiriborde P, Coors A, DeLeo PC, Dyer SD, Ericson JF, Gagne F, Giesy JP, Gouin T, Hallstrom L, Karlsson MV, Larsson DGJ, Lazorchak JM, Mastrocco F, McLaughlin A, McMaster ME, Meyerhoff RD, Moore R, Parrott JL, Snape JR, Murray-Smith R, Servos MR, Sibley PK, Straub JO, Szabo ND, Topp E, Tetreault GR, Trudeau VL, Van der Kraak G. 2012. Pharmaceuticals and personal care products in the environment: What are the big questions? *Environ Health Perspect* 120:1221–1229.

Brack W, Dulio V, Agerstrand M, Allan I, Altenburger R, Brinkmann M, Bunke D, Burgess RM, Cousins I, Escher BI, Hernandez FJ, Hewitt LM, Hilscherova K, Hollender J, Hollert H, Kase R, Klauer B, Lindim C, Lopez Herraez D, Miege C, Munthe J, O'Toole S, Posthuma L, Rudel H, Schafer RB, Sengl M, Smedes F, van de Meent D, van den Brink PJ, van Gils J, van Wezel AP, Vethaak AD, Vermeirssen E, von der Ohe C, Vrana B. 2017. Towards the review of the European Union Water Framework Directive: Recommendations for more efficient assessment and management of chemical contamination in European surface water resources. *Sci Total Environ* 576:720–737.

Braunbeck T, Bottcher M, Hollert H, Kosmehl T, Lammer E, Leist E, Rudolf M, Seitz N. 2005. Towards an alternative for the acute fish LC50 test in

chemical assessment: The fish embryo toxicity test goes multi-species—An update. *ALTEX* 22:87–102.

Brinkmann T, Giner Santonja G, Yükseler H, Roudier S, Delgado Sancho L. 2016. Best available techniques (BAT) reference document for common waste water and waste gas treatment/management systems in the chemical sector. Industrial Emissions Directive 2010/75/EU. European Commission Joint Research Centre, Brussels, Belgium.

Canadian Minister of Justice. 1985. Fisheries act (r.S.C. 1985, c. F-14), section 36. Ottawa, ON.

Castaño A, Bols N, Braunbeck T, Dierickx P, Halder M, Isomaa B, Kawahara K, Lee LEJ, Mothersill C, Part P, Repetto G, Sintes JR, Rufli H, Smith RW, Wood C, Segner H. 2003. The use of fish cells in ecotoxicology: The report and recommendations of ECVAM Workshop 47. *Altern Lab Anim* 31:317–351.

de Vlaming V, Norberg-King T. 1999. A review of single species toxicity tests: Are the tests reliable predictors of aquatic ecosystem responses? US Environmental Protection Agency, Duluth, MN.

Embry MR, Belanger SE, Braunbeck TA, Galay-Burgos M, Halder M, Hinton DE, Leonard MA, Lillicrap A, Norberg-King T, Whale G. 2010. The fish embryo toxicity test as an animal alternative method in hazard and risk assessment and scientific research. *Aquat Toxicol* 97:79–87.

Environment Canada. 2012a. Metal mining environmental effects monitoring (EEM). Gatineau, QC.

Environment Canada. 2012b. Status report on the pulp and paper effluent regulations. Gatineau, QC.

European Commission. 2000. Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for community action in the field of water policy (Water Framework Directive). Official J Eur Commun L327:1–72.

European Commission. 2008. Directive 2008/56/EC of the European Parliament and of the Council of 17 June 2008 establishing a framework for community action in the field of marine environmental policy (Marine Strategy Framework Directive). *Official J Eur Commun* 164:19–40.

European Commission. 2010. Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes. *Official J Eur Union* 276:79.

European Parliament and Council. 2006. Regulation (EC) no. 1907/2006 of the European Parliament and of the Council of 18 December 2006 concerning the registration, evaluation, authorisation and restriction of chemicals (REACH), establishing a European Chemicals Agency, amending Directive 1999/45/EC and repealing Council Regulation (EEC) 793/93 and Commission Regulation (EC) No. 1488/94 as well as Council Directive 76/769/EEC and Commission Directives 91/155/EEC, 93/67/EEC, 93/105/EC and 2000/21/EC. *Official J Eur Union* L396:374–375.

European Union. 2009. Regulation (EC) 1223/2009 of the European Parliament and the Council of 30 November 2009 on cosmetic products. *Official J Eur Union* 342:59–209.

Federal Law Gazette. 2009. Gesetz zur ordnung des wasserhaushalts (wasserhaushaltsgesetz-WHG) (Law on the order of the Water Board). Series 1. German Federal Law Publications, Berlin, Germany.

Federal Law Gazette. 2016. Gesetz ueber abgaben für das einleiten von abwasser in gewaesser (abwasserabgabengesetz-ABWAG). (Wastewater charges act as amended on 18 January 2005); last amended by article 2 of the ordinance of 1 June 2016. Series 1. German Federal Law Publications, Berlin, Germany.

German Institute for Standardisation. 1989. German standard methods for the examination of water, waste water and sludge; bioassays (group l); determining the tolerance of fish to the toxicity of waste water by way of a dilution series (l 31). Din 38412-31:1989–03. Berlin, Germany.

Grothe DR, Dickson KL, Reed-Judkins DK. 1996. *Whole Effluent Toxicity Testing: An Evaluation of Methods and Prediction of Receiving System Impacts*. SETAC, Pensacola, FL, USA.

Halder M, Leonard MA, Iguchi T, Oris JT, Ryder K, Belanger SE, Braunbeck TA, Embry MR, Whale G, Norberg-King TJ, Lillicrap A. 2010. Regulatory aspects on the use of fish embryos in environmental toxicology. *Integr Environ Assess Manag* 6: 484–491.

Hamers T, Legler J, Blaha L, Hylland K, Marigomez I, Schipper CA, Segner H, Vethaak AD, Witters H, de Zwart D, Leonards PE. 2013. Expert opinion on toxicity profiling—Report from a NORMAN expert group meeting. *Integr Environ Assess Manag* 9:185–191.

Hart WB, Doudoroff P, Greenbank J. 1945. *The Evaluation of Toxicity of Industrial Wastes, Chemicals and Other Substances to Freshwater Fishes*. Atlantic Refining. Philadelphia, PA, USA.

Hecker M, Hollert H. 2011. Endocrine disruptor screening: Regulatory perspectives and needs. *Environ Sci Eur* 23:15.

Johnson AC, Sumpter JP. 2001. Removal of endocrine-disrupting chemicals in activated sludge treatment works. *Environ Sci Technol* 35:4697–4703.

Leonards PE, Postma JF, Comber M, Whale G, Stalter G. 2011. Impact of biodegradation on the potential bioaccumulation and toxicity of refinery effluents. *Environ Toxicol Chem* 30:2175–2183.

Lillicrap A, Belanger S, Burden N, Pasquier DD, Embry MR, Halder M, Lampi MA, Lee L, Norberg-King T, Rattner BA, Schirmer K, Thomas P. 2016. Alternative approaches to vertebrate ecotoxicity tests in the 21st century: A review of developments over the last 2 decades and current status. *Environ Toxicol Chem* 35:2637–2646.

Martel PH, O'Connor BI, Kovacs TG, van den Heuvel MR, Parrott JL, McMaster ME, MacLatchy DL, Van Der Kraak GJ, Hewitt LM. 2017. The relationship between organic loading and effects on fish reproduction for pulp mill effluents across Canada. *Environ Sci Technol* 51:3499–3507.

Mills LJ, Chichester C. 2005. Review of evidence: Are endocrine-disrupting chemicals in the aquatic environment impacting fish populations? *Sci Total Environ* 343:1–34.

National Centre for the Replacement Refinement & Reduction of Animals in Research. 2017. Centre for the Replacement Refinement & Reduction of Animals in Research. 2017. Challenge 10: Predart. [cited the authors January 9]. https://crackit.org.uk/challenge-10-predart

Nonet S. 2005. Requirements and test methods for on site domestic wastewater treatment plants: The European standard (prEN 12566-3) compared to other international standards. *Water Sci Technol* 51:147–154.

Norberg-King TJ, Ausley LW, Burton DT, Goodfellow WL, Miller JL, Waller WT. 2005. *Toxicity Reduction and Toxicity Identification Evaluations (TIE) for Effluents, Ambient Waters, and Other Aqueous Media*. SETAC, Pensacola, FL, USA.

Organisation for Economic Co-operation and Development. 2013. Test No. 236: Fish embryo acute toxicity (FET) test. *Guidelines for the testing of chemicals*. Paris, France.

OSPAR Commission. 2000. OSPAR background document concerning the elaboration of programmes and measures relating to whole effluent assessment. London, UK.

OSPAR Commission. 2007. Practical guidance document on whole effluent assessment. Publication 316/2007. London, UK.

Parkerton TF, Stone MA, Letinski DJ. 2000. Assessing the aquatic toxicity of complex hydrocarbon mixtures using solid phase microextraction. *Toxicol Lett* 112:273–282.

Pelka KE, Henn K, Keck A, Sapel B, Braunbeck T. 2017. Size does matter—Determination of the critical molecular size for the uptake of chemicals across the chorion of zebrafish (*Danio rerio*) embryos. *Aquat Toxicol* 185:1–10.

Power EA, Boumphrey RS. 2004. International trends in bioassay use for effluent management. *Ecotoxicology* 13:377–398.

Racz PI, Wildwater M, Rooseboom M, Kerkhof E, Pieters R, Yebra-Pimentel ES, Dirks RP, Spaink HP, Smulders C, Whale GF. 2017. Application of *Caenorhabditis elegans* (nematode) and *Danio rerio* embryo (zebrafish) as model systems to screen for developmental and reproductive toxicity of piperazine compounds. *Toxicol In Vitro* 44:11–16.

Russell WMS, Burch SL. 1959. *The Principles of Humane Experimental Technique*. Methuen, London, UK.

Schirmer K. 2006. Proposal to improve vertebrate cell cultures to establish them as substitutes for the regulatory testing of chemicals and effluents using fish. *Toxicology* 224:163–183.

Scholz S, Renner P, Belanger SE, Busquet F, Davi R, Demeneix BA, Denny JS, Leonard M, McMaster ME, Villeneuve DL, Embry MR. 2013. Alternatives to in vivo tests to detect endocrine disrupting chemicals (EDCs) in fish and amphibians—Screening for estrogen, androgen and thyroid hormone disruption. *Crit Rev Toxicol* 43:45–72.

Segner H. 2011. Moving beyond a descriptive aquatic toxicology: The value of biological process and trait information. *Aquat Toxicol* 105(3–4 Suppl.):50–55.

UK Environment Agency. 2006. Performance standard for laboratories undertaking direct toxicity assessment of effluents. London, UK.

US Environmental Protection Agency. 2002. Guidelines establishing test procedures for the analysis of pollutants; whole effluent toxicity test methods; final rule. *Fed Reg* 67:69951–69972.

US Environmental Protection Agency. 1991. Technical support document for water quality–based toxics control. Washington, DC.

US Environmental Protection Agency. 2016. The Frank R. Lautenberg Chemical Safety for the 21st Century Act. [cited 2018 January 9]. Available from: https://www.epa.gov/assessing-and-managing-chemicals-under-tsca/frank-r-lautenberg-chemical-safety-21st-century-act

Volz DC, Belanger S, Embry M, Padilla S, Sanderson H, Schirmer K, Scholz S, Villeneuve D. 2011. Adverse outcome pathways during early fish development: A conceptual framework for identification of chemical screening and prioritization strategies. *Toxicol Sci* 123:349–358.

Warren CE. 1971. *Biology and Water Pollution Control*. W.B. Saunders, Philadelphia, PA.

Whale GF, Quill SN, Eadsforth CV. 2003. Evaluating alternatives to the use of fish for environmental assessments. *Proceedings*, SETAC Europe 13th Annual Meeting, Hamburg, Germany, April 27-May 1, 2003.

